統計数理(2005)
第 53 巻 第 2 号 331–348

特集「計算推論—モデリング・数理・アルゴリズム—」
［原著論文］

GMRES 法による最小二乗問題の解法

速水 謙[1,2]・伊藤 徳史[3]

（受付 2005 年 3 月 24 日；改訂 2005 年 9 月 20 日）

要　　旨

大規模疎な $m \times n$ 行列 A を係数行列とする最小二乗問題に対する主流の反復法は，正規方程式に対して（前処理付き）共役勾配法を適用する CGLS 法である．

本論文ではまず，代案として元の最小二乗問題を，$n \times m$ 行列 B を用いて，正方行列 AB または BA を係数行列にもつ等価な最小二乗問題に変形し，非対称正方行列を係数行列とする連立一次方程式用のロバストなクリロフ部分空間反復法である一般化最小残差法（GMRES）を適用する手法を提案する．

次に，$m \geq n$（優決定），$m < n$（劣決定），およびランク落ちも含めた一般の場合に対して，これらの手法が任意の右辺項 b に対して破綻することなく最小二乗解を与えるための行列 B に関する十分条件を導く．そして，B の例として不完全 QR 分解の一つである IMGS(l) 法を提案する．

最後に，フルランクな優決定および劣決定問題に対する数値実験により，条件の悪い問題では（対角スケーリングに相当する）IMGS(0) 法を用いた提案手法の方が従来の前処理付き CGLS 法より速く最小二乗解を与えることを示す．

キーワード：最小二乗問題，クリロフ部分空間反復法，CGLS 法，GMRES 法，特異な系，前処理．

1. はじめに

最小二乗問題

$$\min_{\boldsymbol{x} \in \mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2 \tag{1.1}$$

を考える．ただし，A は $m \times n$ の実行列で，$m \geq n$ と $m < n$ の両方の場合を考えるものとする．また，$\mathrm{rank} A \leq \min(m, n)$ の等号が成り立たない，いわゆるランク落ちの場合も許すものとする．

式(1.1)は正規方程式

$$A^{\mathrm{T}} A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b} \tag{1.2}$$

と等価である．

[1] 国立情報学研究所 情報学基礎研究系：〒101–8430 東京都千代田区一ツ橋 2–1–2; hayami@nii.ac.jp
[2] 総合研究大学院大学 複合科学研究科情報学専攻
[3] 株式会社 ビジネスデザイン研究所：〒100–0006 東京都千代田区有楽町 1–7–1 有楽町電気ビル南館 9F; ito@business-design.co.jp

1.1 直接法

上記の問題(で $m \geq n$ でランク落ちのない場合)に対する従来法としては，まず直接法がある．代表的な方法としては，まず A を，$m \times n$ の直交行列 $Q\,(Q^{\mathrm{T}}Q = \mathrm{I}_n)$ と，$n \times n$ の上三角行列 R の積に $A = QR$ と分解する．具体的には，Householder 法，修正 Gram-Schmidt 法，Givens 法などを用いる．すると，(1.2)式は $R^{\mathrm{T}}R\boldsymbol{x} = R^{\mathrm{T}}Q^{\mathrm{T}}\boldsymbol{b}$ と等価である．また，$\mathrm{rank}A = n$ なら R は正則だから，$R\boldsymbol{x} = Q^{\mathrm{T}}\boldsymbol{b}$ から後退代入により，式(1.1)の最小二乗解 x が求まる．

ただし，A が大規模で疎な場合は，メモリーと計算時間を節約する工夫が必要となる(Björck, 1996)．

1.2 正規方程式を用いた反復法

大規模疎な問題に対しては，メモリーと計算時間の節約のため反復法が有効となる．反復法による最小二乗問題の従来の解法に関しては例えば Björck (1996)を参照されたい．なお，本研究に関連する先駆的な研究として Tanabe (1974, 1975)がある．

正規方程式(1.2)の係数行列 $A^{\mathrm{T}}A$ は対称な正方行列であり，$\mathrm{rank}A = n$ なら定値なので，正規方程式に共役勾配法を適用する下記の Conjugate Gradient Least Squares (CGLS)法が従来主流の方法である．

CGLS 法

Choose $\boldsymbol{x}_0$.
$\tilde{\boldsymbol{r}}_0 = A^{\mathrm{T}}(\boldsymbol{b} - A\boldsymbol{x}_0)$
$\boldsymbol{p}_0 = \tilde{\boldsymbol{r}}_0$
for $i = 0, 1, 2, \ldots$ until $\|\tilde{\boldsymbol{r}}_i\|_2 < \epsilon$
$$\alpha_i = \frac{(\tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{r}}_i)}{(\boldsymbol{p}_i, A^{\mathrm{T}}A\boldsymbol{p}_i)}$$
$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i \boldsymbol{p}_i$$
$$\tilde{\boldsymbol{r}}_{i+1} = \tilde{\boldsymbol{r}}_i - \alpha_i A^{\mathrm{T}}A\boldsymbol{p}_i$$
$$\beta_i = \frac{(\tilde{\boldsymbol{r}}_{i+1}, \tilde{\boldsymbol{r}}_{i+1})}{(\tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{r}}_i)}$$
$$\boldsymbol{p}_{i+1} = \tilde{\boldsymbol{r}}_{i+1} + \beta_i \boldsymbol{p}_i$$
endfor

ところで，一般に長方形行列 A の一般逆行列を $A^{\dagger}$，$\mathrm{rank}A = r$，最大，最小特異値を各々 σ_1, σ_r とすると，A の条件数は

$$\kappa(A) := \|A\|_2 \, \|A^{\dagger}\|_2 = \frac{\sigma_1}{\sigma_r}$$

で与えられる(例えば Björck, 1996 参照)．よって，

$$\kappa(A^{\mathrm{T}}A) = \left(\frac{\sigma_1}{\sigma_r}\right)^2 = \kappa(A)^2$$

となる．

CGLS 法の収束の速さは(A がランク落ちしている場合も含めて)A の条件数 $\kappa(A)$ に依存することが知られている(Björck, 1996 参照)．従って，悪条件問題に対しては CGLS 法そのものの収束は遅いので，適切な前処理が必要となる．簡便な前処理法としては，$A^{\mathrm{T}}A$ の対角項を用いたスケーリングがある．その他に，例えば不完全コレスキー分解(Meijerink and van der Vorst, 1977)，不完全 QR 分解(Jennings and Ajiz, 1984; Saad, 1988)，不完全 Givens 直交化(Bai et al., 2001)，ロバストな不完全分解(Benzi and Tuma, 2003)などの手法が提案されている．

例えば不完全 QR 分解の例として Jennings and Ajiz (1984) による下記の不完全修正 Gram-Schmidt (IMGS) 法がある．

$A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$, $\boldsymbol{a}_i^{(1)} = \boldsymbol{a}_i \quad (i = 1, \ldots, n), \tau$ を閾値パラメタとして，

for $i = 1, 2, \ldots, n$

$\quad r_{ii} = \|\boldsymbol{a}_i^{(i)}\|_2, \qquad \boldsymbol{q}_i = \dfrac{\boldsymbol{a}_i^{(i)}}{r_{ii}}$

$\quad$ for $j = i+1, \ldots, n$

$\qquad r_{ij} = \boldsymbol{q}_i^{\mathrm{T}} \boldsymbol{a}_j^{(i)}$

$\qquad$ if $|r_{ij}| < \tau \|\boldsymbol{a}_i\|_2$ then $r_{ij} = 0$

$\qquad \boldsymbol{a}_j^{(i+1)} = \boldsymbol{a}_j^{(i)} - r_{ij}\boldsymbol{q}_i$

$\quad$ endfor

endfor

また，Saad (1988) は下記のような不完全修正 Gram-Schmidt (IMGS) 法を提案している．

p_Q, p_R をパラメタとして，

for $i = 1, 2, \ldots, n$

$\quad r_{ii} = \|\boldsymbol{a}_i^{(i)}\|_2, \qquad \boldsymbol{q}_i = \dfrac{\boldsymbol{a}_i^{(i)}}{r_{ii}}$

$\quad$ Determine the p_Q largest elements of $\boldsymbol{q}_i$ and assign 0 to the other elements.

$\quad$ for $j = i+1, \ldots, n$

$\qquad r_{ij} = \boldsymbol{q}_i^{\mathrm{T}} \boldsymbol{a}_j^{(i)}$

$\qquad$ Determine the p_R largest r_{ij}'s for $i+1 \le j \le n$ and assign 0 to the others.

$\qquad \boldsymbol{a}_j^{(i+1)} = \boldsymbol{a}_j^{(i)} - r_{ij}\boldsymbol{q}_i$

$\quad$ endfor

endfor

このようにして得られた上三角行列 $R = (r_{ij})$ を用いて正規方程式 (1.2) を

$$\tilde{A}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}}, \tag{1.3}$$

ただし，

$$\tilde{A} = R^{-\mathrm{T}} A^{\mathrm{T}} A R^{-1}, \qquad \tilde{\boldsymbol{x}} = R\boldsymbol{x}, \qquad \tilde{\boldsymbol{b}} = R^{-\mathrm{T}} A^{\mathrm{T}} \boldsymbol{b},$$

と前処理し，式 (1.3) に CGLS 法を適用する．

1.3 正規方程式に基づかない反復法

前処理をしたとしても，式 (1.3) の条件数は AR^{-1} の条件数の二乗なので悪条件問題に対しては収束がまだ十分速くない可能性がある．これに対して，張 (1989)，Zhang and Oyanagi (1990, 1991) は正規方程式に基づかないで，A を直接扱う反復法として，$n \times m$ の写像行列 B を用いて行列 AB を係数行列とする系に対して，非対称正方行列を係数とする連立一次方程式のクリロフ部分空間反復解法の一つである Orthomin(k) 法を適用する，CR-LS(k) 法を提案している．

本論文では，Orthomin(k) 法よりも破綻しにくい GMRES 法に同様の考えを適用することから始める (伊藤・速水, 2002, 2003; 伊藤, 2003)．次に，行列 BA を係数行列とする系に対して GMRES 法を適用する方法を提案する (Ito and Hayami, 2004; 伊藤・速水, 2004; 速水・伊藤, 2004, 2005)．そして，$m \ge n$ の場合だけでなく，$m < n$ の場合や，A がフルランクでない場合も含めて，両手法が任意の b に対して破綻することなく最小二乗解を与えるための行列 B に関する十分条件を導く．最後に，フルランクな優決定および劣決定問題に対する数値実験によ

り両手法の有効性を示す．

2. GMRES 法による最小二乗問題の解法

Saad and Schultz (1986) の GMRES 法 (Generalized Minimal Residual method) は，対称とは限らない正則な行列を係数行列とする連立一次方程式のロバストなクリロフ部分空間法として知られる．そのままだと，反復数の増加に伴いメモリーと計算量が膨大になるので，下記のように k 反復毎にそのときの近似解を初期解としてアルゴリズムを再開する（リスタートする）GMRES(k) 法を使うことが多い．($k=\infty$ としたのが GMRES 法に相当する．)

GMRES(k) 法

Choose $\boldsymbol{x}_0$.

$*$ $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$

$$\boldsymbol{v}_1 = \frac{\boldsymbol{r}_0}{\|\boldsymbol{r}_0\|_2}$$

for $i = 1, 2, \ldots, k$

$h_{j,i} = (A\boldsymbol{v}_i, \boldsymbol{v}_j) \qquad (j = 1, 2, \ldots, i)$

$$\hat{\boldsymbol{v}}_{i+1} = A\boldsymbol{v}_i - \sum_{j=1}^{i} h_{j,i}\boldsymbol{v}_j$$

$h_{i+1,i} = \|\hat{\boldsymbol{v}}_{i+1}\|_2$

$$\boldsymbol{v}_{i+1} = \frac{\hat{\boldsymbol{v}}_{i+1}}{h_{i+1,i}}$$

Find $\boldsymbol{y}_i \in \mathbf{R}^i$ which minimizes $\|\boldsymbol{r}_i\|_2 = \| \ \|\boldsymbol{r}_0\|_2 \boldsymbol{e}_i - \bar{H}_i \boldsymbol{y}\|_2$.

if $\|\boldsymbol{r}_i\|_2 < \epsilon$ then

$\boldsymbol{x}_i = \boldsymbol{x}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i$

stop

endif

endfor

$\boldsymbol{x}_0 = \boldsymbol{x}_k$

Go to $*$.

ただし，上記で $\bar{H}_i = (h_{pq}) \in \mathbf{R}^{(i+1)\times i}, \quad \boldsymbol{e}_i = (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbf{R}^{i+1}$ とする．

上記で $h_{i+1,i} = 0$ となるとき，アルゴリズムが破綻するという．丸め誤差がなければ，$A \in \mathbf{R}^{N\times N}$ が正則な場合は GMRES 法は破綻することなく N 反復以内に連立一次方程式 $Ax = \boldsymbol{b}$ の真の解を与える (Saad and Schultz, 1986)．A が特異な場合の GMRES 法の振る舞いについては，例えば Brown and Walker (1997) や 速水 (2003) で論じられている．

さて，GMRES 法を直接式 (1.1) の最小二乗問題に適用しようとしても，A は $m \times n$ 行列で，初期近似解 x_0 に対する残差ベクトル $r_0 = \boldsymbol{b} - A\boldsymbol{x}_0$ は m 次元ベクトルなので，r_0 に A をかけてクリロフ部分空間を構成することはできない．そこで，$n \times m$ 行列 B を使ってこの問題を解決するには二つの方法が考えられる．

2.1 方法 1

まず最初の方法は張 (1989)，Zhang and Oyanagi (1990, 1991) のように A の右から B をかけて $m \times m$ 行列 AB を構成し，m 次元空間の中のクリロフ部分空間 $\langle r_0, AB\boldsymbol{r}_0, \ldots, (AB)^{i-1}\boldsymbol{r}_0 \rangle$ を用いるものである．

ここで，行列 M に対して $\mathcal{R}(M)$ を M の像空間，$\mathcal{N}(M)$ を M の核空間，部分空間 V に対して $V^{\perp}$ を V の直交補空間として以下の準備をする．

定理 2.1. 任意の $\boldsymbol{b}\in\mathbf{R}^m$ に対して $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2=\min_{\boldsymbol{z}\in\mathbf{R}^m}\|\boldsymbol{b}-AB\boldsymbol{z}\|_2$ が成り立つための必要十分条件は $\mathcal{R}(A)=\mathcal{R}(AB)$ である．

証明.($\Longleftarrow$): 自明.
($\Longrightarrow$):

$$\mathcal{R}(A)\neq\mathcal{R}(AB)\Longrightarrow\mathcal{R}(A)\supset\mathcal{R}(AB)\Longrightarrow{}^{\exists}\tilde{\boldsymbol{b}}\in\mathcal{R}(A)\backslash\mathcal{R}(AB)$$

$$\Longleftrightarrow{}^{\exists}\tilde{\boldsymbol{x}}\in\mathbf{R}^n;\quad \tilde{\boldsymbol{b}}=A\tilde{\boldsymbol{x}},\quad \text{しかし}\tilde{\boldsymbol{b}}\neq AB\boldsymbol{z}\quad {}^{\forall}\boldsymbol{z}\in\mathbf{R}^m$$

$$\Longleftrightarrow 0=\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\tilde{\boldsymbol{b}}-A\boldsymbol{x}\|_2<\min_{\boldsymbol{z}\in\mathbf{R}^m}\|\tilde{\boldsymbol{b}}-AB\boldsymbol{z}\|_2. \qquad \square$$

補題 2.2. $\mathcal{R}(AA^{\mathrm{T}})=\mathcal{R}(A)$.

証明.

$$\begin{aligned}\mathcal{R}(AA^{\mathrm{T}})&=\left\{A\boldsymbol{x}\,|\,\boldsymbol{x}\in\mathcal{R}(A^{\mathrm{T}})\right\}=\left\{A\boldsymbol{x}\,|\,\boldsymbol{x}\in\mathcal{N}(A)^{\perp}\right\}\\&=\left\{A\boldsymbol{x}\,|\,\boldsymbol{x}\in\mathcal{N}(A)^{\perp}\cup\mathcal{N}(A)\right\}=\mathcal{R}(A).\end{aligned} \qquad \square$$

（例えば，ラオ, 1977, p. 26 の別証も参照.）

補題 2.3. $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)\Longrightarrow\mathcal{R}(A)=\mathcal{R}(AB)$.

証明．補題 2.2 より明らか． $\square$

例えば，$\mathrm{rank}A=\mathrm{rank}B=n$ ならば $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)=\mathbf{R}^n$ が成り立つ．
そこで $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$ を仮定すると，補題 2.3 より $\mathcal{R}(A)=\mathcal{R}(AB)$ が成り立ち，任意の $x_0\in\mathbf{R}^n$ に対して $Ax_0=ABz_0$ となるような $z_0\in\mathbf{R}^m$ が存在し，$\boldsymbol{r}_0=\boldsymbol{b}-A\boldsymbol{x}_0=\boldsymbol{r}_0-AB\boldsymbol{z}_0$ となる．
そこで GMRES(k)法を使って最小二乗問題

$$\min_{\boldsymbol{z}\in\mathbf{R}^m}\|\boldsymbol{b}-AB\boldsymbol{z}\|_2=\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$$

を解くことを考える．その際，初期近似解 $z=z_0$ を $ABz_0=Ax_0$ となるようにとると，下記のアルゴリズムを得る．

GMRES(k)-LS 法 1

Choose $\boldsymbol{x}_0\quad(A\boldsymbol{x}_0=AB\boldsymbol{z}_0)$.
* $\boldsymbol{r}_0=\boldsymbol{b}-A\boldsymbol{x}_0\quad(=\boldsymbol{b}-AB\boldsymbol{z}_0)$
$\boldsymbol{v}_1=\dfrac{\boldsymbol{r}_0}{\|\boldsymbol{r}_0\|_2}$

for $i = 1, 2, \ldots, k$
$h_{j,i} = (AB\boldsymbol{v}_i, \boldsymbol{v}_j) \qquad (j = 1, 2, \ldots, i)$
$$\hat{\boldsymbol{v}}_{i+1} = AB\boldsymbol{v}_i - \sum_{j=1}^{i} h_{j,i}\boldsymbol{v}_j$$
$h_{i+1,i} = \|\hat{\boldsymbol{v}}_{i+1}\|_2$
$$\boldsymbol{v}_{i+1} = \frac{\hat{\boldsymbol{v}}_{i+1}}{h_{i+1,i}}$$
Find $\boldsymbol{y}_i \in \mathbf{R}^i$ which minimizes $\|\boldsymbol{r}_i\|_2 = \| \ \|\boldsymbol{r}_0\|_2 \, \boldsymbol{e}_i - \bar{H}_i \, \boldsymbol{y} \|_2$.
$\boldsymbol{x}_i = \boldsymbol{x}_0 + B[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i \qquad (\Longleftrightarrow \boldsymbol{z}_i = \boldsymbol{z}_0 + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i]\boldsymbol{y}_i)$
$\boldsymbol{r}_i = \boldsymbol{b} - A\boldsymbol{x}_i$
if $\|A^{\mathrm{T}}\boldsymbol{r}_i\|_2 < \epsilon$ stop
endfor
$\boldsymbol{x}_0 = \boldsymbol{x}_k \qquad (\Longleftrightarrow \boldsymbol{z}_0 = \boldsymbol{z}_k)$
Go to $*$.

ところで，上記のアルゴリズムで GMRES(k)法の代わりに GMRES 法を用いた方法($k = \infty$ の場合)を GMRES-LS 法 1 と呼ぶことにする．この GMRES-LS 法 1 が破綻することなく式(1.1)の最小二乗解を与えるための必要十分条件を考える．

まず，下記の定理(Brown and Walker, 1997; 速水, 2003)に注意する．

定理 2.4. $\tilde{A} \in \mathbf{R}^{m \times m}$ に対して下記が成り立つ．

GMRES 法が任意の $\boldsymbol{b} \in \mathbf{R}^m$ および任意の初期近似解 $z_0 \in \mathbf{R}^m$ に対して破綻することなく $\min_{\boldsymbol{z} \in \mathbf{R}^m} \|\boldsymbol{b} - \tilde{A}\boldsymbol{z}\|_2$ の最小二乗解を与えるための必要十分条件は $\mathcal{R}(\tilde{A}) = \mathcal{R}(\tilde{A}^{\mathrm{T}})$ である．

次に下記が成り立つ．

定理 2.5. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ ならば，「$\mathcal{R}(AB) = \mathcal{R}(B^{\mathrm{T}}A^{\mathrm{T}}) \Longleftrightarrow \mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ 」が成り立つ．

証明. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ ならば，補題 2.2 より

$$\mathcal{R}(AB) = \mathcal{R}(AA^{\mathrm{T}}) = \mathcal{R}(A), \qquad \mathcal{R}(B^{\mathrm{T}}A^{\mathrm{T}}) = \mathcal{R}(B^{\mathrm{T}}B) = \mathcal{R}(B^{\mathrm{T}}). \qquad \square$$

従って，定理 2.4 で $\tilde{A} = AB$ とおくと，定理 2.5 より，下記を得る．

定理 2.6. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ ならば下記が成り立つ．

GMRES-LS 法 1 が任意の $\boldsymbol{b} \in \mathbf{R}^m$ および任意の $x_0 \in \mathbf{R}^n$ に対して破綻することなく $\min_{\boldsymbol{x} \in \mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ の最小二乗解を与えるための必要十分条件は $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ である．

また，系として下記を得る．

系 2.7. $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ かつ $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}})$ ならば GMRES-LS 法 1 は任意の $\boldsymbol{b} \in \mathbf{R}^m$ および任意の $x_0 \in \mathbf{R}^n$ に対して破綻することなく $\min_{\boldsymbol{x} \in \mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$ の最小二乗解を与える．

ここで，$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ は，B がある正則な行列 C を用いて

$$B=CA^{\mathrm{T}} \tag{2.1}$$

と表されることと等価である．

ところで，Calvetti et al.（2000），Reichel and Ye（2005）は GMRES 法を用いて $m \geq n$ の場合の最小二乗問題を解く方法として，行列 A の右側に 0 列ベクトルを $m-n$ 本付け足して特異行列 $\tilde{A}=[A,0]$ を構成し，GMRES 法を

$$\min_{\boldsymbol{z}\in\mathbf{R}^m}\|\boldsymbol{b}-\tilde{A}\boldsymbol{z}\|_2 \ \left(=\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2\right)$$

に適用させることを提案している．

これは，上記の GMRES-LS 法 1 において $B=[\mathrm{I}_n,0]$，$\tilde{A}=AB$ とおくことと等価である．ただし，I_n は $n\times n$ の単位行列である．

従って，$\mathrm{rank}A=n$ ならば $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)=\mathbf{R}^n$ は成立するが，$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ は成立するとは限らないので，定理 2.6 より彼らの方法は最小二乗解に到達する前に破綻する可能性がある．

2.2 方法 2

もう一つの方法として，同じ $n\times m$ の写像行列 B を用いて下記のように初期残差ベクトル $\boldsymbol{r}_0\in\mathbf{R}^m$ を $\tilde{\boldsymbol{r}}_0=B\boldsymbol{r}_0\in\mathbf{R}^n$ に写像しておいて，n 次元空間内のクリロフ部分空間 $\langle\tilde{\boldsymbol{r}}_0, BA\tilde{\boldsymbol{r}}_0,\ldots,(BA)^{i-1}\tilde{\boldsymbol{r}}_0\rangle$ を構成し GMRES(k)法を適用する次の方法も考えられる．

GMRES(k)-LS 法 2

Choose $\boldsymbol{x}_0$.
$*$ $\tilde{\boldsymbol{r}}_0=B(\boldsymbol{b}-A\boldsymbol{x}_0)$
$\boldsymbol{v}_1=\tilde{\boldsymbol{r}}_0/\|\tilde{\boldsymbol{r}}_0\|_2$
for $i=1,2,\ldots,k$
  $h_{j,i}=(BA\boldsymbol{v}_i,\boldsymbol{v}_j)\quad(j=1,2,\ldots,i)$
  $\hat{\boldsymbol{v}}_{i+1}=BA\boldsymbol{v}_i-\sum_{j=1}^{i}h_{j,i}\boldsymbol{v}_j$
  $h_{i+1,i}=\|\hat{\boldsymbol{v}}_{i+1}\|_2$
  $\boldsymbol{v}_{i+1}=\hat{\boldsymbol{v}}_{i+1}/h_{i+1,i}$
  Find $\boldsymbol{y}_i\in\mathbf{R}^i$ which minimizes $\|\tilde{\boldsymbol{r}}_i\|_2=\|\ \|\tilde{\boldsymbol{r}}_0\|_2\boldsymbol{e}_i-\bar{H}_i\boldsymbol{y}\|_2$.
  $\boldsymbol{x}_i=\boldsymbol{x}_0+[\boldsymbol{v}_1,\ldots,\boldsymbol{v}_i]\boldsymbol{y}_i$
  $\boldsymbol{r}_i=\boldsymbol{b}-A\boldsymbol{x}_i$
  if $\|A^{\mathrm{T}}\boldsymbol{r}_i\|_2<\epsilon$ stop
endfor
$\boldsymbol{x}_0=\boldsymbol{x}_k$
Go to $*$.

この手法は

$$BA\boldsymbol{x}=B\boldsymbol{b} \tag{2.2}$$

に GMRES(k)法を適用するのと等価である．

ここでまず下記が成り立つ．（この定理および証明は査読者によって大幅に改善された結果である．）

定理 2.8. 任意の $\boldsymbol{b}\in\mathbf{R}^m$ に対して方程式 $BA\boldsymbol{x}=B\boldsymbol{b}$ が解をもち，その解が $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の解となるための必要十分条件は $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ である．

証明．($\Longrightarrow$) まず，任意の $\boldsymbol{b}\in\mathbf{R}^m$ に対して方程式 $BA\boldsymbol{x}=B\boldsymbol{b}$ が解をもつための必要十分条件は

$$\mathcal{R}(BA)=\mathcal{R}(B). \tag{2.3}$$

つぎに，方程式 $BA\boldsymbol{x}=B\boldsymbol{b}$ が解をもつとして，その解が $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の解となる，すなわち，任意の $\boldsymbol{b}\in\mathbf{R}^m$ に対して，$B(\boldsymbol{b}-A\boldsymbol{x})=\mathbf{0}$ ならば，$A^{\mathrm{T}}(\boldsymbol{b}-A\boldsymbol{x})=\mathbf{0}$ となるための必要十分条件は

$$\mathcal{N}(B)\subseteq\mathcal{N}(A^{\mathrm{T}}).$$

これを像空間で書けば，

$$\mathcal{R}(A)\subseteq\mathcal{R}(B^{\mathrm{T}}). \tag{2.4}$$

ここで

$$\mathrm{rank}BA\le\min\{\mathrm{rank}A,\mathrm{rank}B\},$$

すなわち

$$\dim\mathcal{R}(BA)\le\min\{\dim\mathcal{R}(B),\dim\mathcal{R}(A)\}$$

に注意すれば，(2.3) より，$\dim\mathcal{R}(B)\le\dim\mathcal{R}(A)$ を得る．これは，$\dim\mathcal{R}(B^{\mathrm{T}})\le\dim\mathcal{R}(A)$ と同値であり，(2.4) と合わせて，$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ を得る．

($\Longleftarrow$)

$$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(BA)=\mathcal{R}(BB^{\mathrm{T}})=\mathcal{R}(B).$$

また，

$$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})\Longrightarrow\mathcal{R}(A)\subseteq\mathcal{R}(B^{\mathrm{T}}).$$ □

次に，定理 2.5 と双対的に下記が成立する．

定理 2.9. $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ ならば，「$\mathcal{R}(BA)=\mathcal{R}(A^{\mathrm{T}}B^{\mathrm{T}})\Longleftrightarrow\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$」が成り立つ．

証明． $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ ならば，補題 2.2 より

$$\mathcal{R}(BA)=\mathcal{R}(BB^{\mathrm{T}})=\mathcal{R}(B),\qquad\mathcal{R}(A^{\mathrm{T}}B^{\mathrm{T}})=\mathcal{R}(A^{\mathrm{T}}A)=\mathcal{R}(A^{\mathrm{T}}).$$ □

さらに，定理 2.4 で $\tilde{A}=BA$ とおき，定理 2.9 を用いると下記を得る．

定理 2.10. $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ ならば下記が成り立つ．
GMRES-LS 法 2 が任意の $\boldsymbol{b}\in\mathbf{R}^m$ および任意の $\boldsymbol{x}_0\in\mathbf{R}^n$ に対して破綻することなく $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の最小二乗解を与えるための必要十分条件は $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$ である．

また，下記の系が成り立つ．

系 2.11. $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ かつ $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$ ならば GMRES-LS 法 2 は任意の $\boldsymbol{b}\in\mathbf{R}^m$ および任意の $x_0\in\mathbf{R}^n$ に対して破綻することなく $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の最小二乗解を与える．

ところで，Reichel and Ye (2005) は GMRES 法を用いて $m\le n$ の場合の最小二乗問題を解く方法として，行列 A の下に 0 行ベクトルを $n-m$ 本付け足して特異行列

$$\tilde{A}=\begin{bmatrix}A\\0\end{bmatrix}$$

を構成し，さらに

$$\tilde{\boldsymbol{b}}=\begin{bmatrix}\boldsymbol{b}\\0\end{bmatrix}$$

として，GMRES 法を

$$\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\tilde{\boldsymbol{b}}-\tilde{A}\boldsymbol{x}\|_2$$

に適用させることを提案している．

これは，上記の GMRES-LS 法 2 において，I_m を $m\times m$ の単位行列として，

$$B=\begin{bmatrix}\mathrm{I}_m\\0\end{bmatrix}$$

とおいて，$\tilde{A}=BA,\quad \tilde{\boldsymbol{b}}=B\boldsymbol{b}$ とおくことと等価である．

従って，$\mathrm{rank}A=m$ ならば $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})=\mathbf{R}^m$ は成立するが，$\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$ は成立するとは限らないので，定理 2.10 より彼らの方法は最小二乗解に到達する前に破綻する可能性がある．

2.3 B の充たすべき条件

以上をまとめると，ランク落ちの場合も含めた $\mathrm{rank}A\le\min(m,n)$ の一般の場合には下記が成り立つ．

条件：

$$\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}}),\qquad \mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B) \tag{2.5}$$

が成り立てば，系 2.7 および系 2.11 より，GMRES-LS 法 1 および 2 は任意の $\boldsymbol{b}\in\mathbf{R}^m$ および任意の $x_0\in\mathbf{R}^n$ に対して $\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の最小二乗解を与える．ここで，条件(2.5)は $B=A^{\mathrm{T}}$ とおけば成り立つ．

次に，$\mathrm{rank}A=\min(m,n)$ のフルランクの場合を考える．まず，$m\ge n=\mathrm{rank}A$ （優決定問題）の場合を考える．任意の正則な行列 $C\in\mathbf{R}^{n\times n}$ を用いて $B=CA^{\mathrm{T}}$ とおくことを考える．$B,A^{\mathrm{T}}\in\mathbf{R}^{n\times m}$ に対して以下が成り立つ．

$$\begin{gathered}B=CA^{\mathrm{T}},\quad C\in\mathbf{R}^{n\times n}:\text{正則}\\ \Updownarrow\\ B^{\mathrm{T}}=AC^{\mathrm{T}},\quad C^{\mathrm{T}}:\text{正則}\\ \Updownarrow\\ \mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}}).\end{gathered}$$

従って，$n = \mathrm{rank}A^{\mathrm{T}} = \mathrm{rank}A = \mathrm{rank}B^{\mathrm{T}} = \mathrm{rank}B$ より $\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B) = \mathbf{R}^n$ も成り立つ．

ここで，$AB \in \mathbf{R}^{m\times m}$, $BA \in \mathbf{R}^{n\times n}$, $m \geq n$ だから，GMRES-LS 法 2（$BA\boldsymbol{x} = B\boldsymbol{b}$）を用いる方が GMRES-LS 法 1 を用いるより反復当たりの計算量が少なくてすむ．$\mathrm{rank}A = n$ ならば $A^{\mathrm{T}}A$，そして $\mathrm{diag}(A^{\mathrm{T}}A)$ は正則なので，C の簡単な例として Zhang and Oyanagi（1991）のように $C := \{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}$ とおくことが考えられる．

次に，$\mathrm{rank}A = m \leq n$（劣決定）の場合は，任意の正則な行列 $C \in \mathbf{R}^{m\times m}$ を用いて $B = A^{\mathrm{T}}C$ とおくことを考える．$B, A^{\mathrm{T}} \in \mathbf{R}^{n\times m}$ に対して以下が成り立つ．

$$B = A^{\mathrm{T}}C, \quad C \in \mathbf{R}^{m\times m} : \text{正則}$$
$$\Updownarrow$$
$$\mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B).$$

従って，$m = \mathrm{rank}A = \mathrm{rank}A^{\mathrm{T}} = \mathrm{rank}B = \mathrm{rank}B^{\mathrm{T}}$ より $\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}) = \mathbf{R}^m$ も成り立つ．

ここで，$AB \in \mathbf{R}^{m\times m}$, $BA \in \mathbf{R}^{n\times n}$, $m \leq n$ だから，GMRES-LS 法 1（$\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$）を用いる方が GMRES-LS 法 2 を用いるより反復当たりの計算量が少なくてすむ．$\mathrm{rank}A = m$ ならば AA^{T}，そして $\mathrm{diag}(AA^{\mathrm{T}})$ は正則なので，C の簡単な例として $C := \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ が考えられる．

3. B の選び方

$\mathcal{R}(A) = \mathcal{R}(B^{\mathrm{T}}), \mathcal{R}(A^{\mathrm{T}}) = \mathcal{R}(B)$ を充たす以外に，収束を速めるには，B は $AB \approx \mathrm{I}_m$ または $BA \approx \mathrm{I}_n$ を充たすことが望ましい．

簡単な候補としては，先にも述べたように，$m \geq n = \mathrm{rank}A$ の場合は $C := \{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}$ とおき，$\mathrm{rank}A = m \leq n$ の場合は $C := \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ とおくことがまず考えられる．

3.1 不完全 QR 分解

ここでは，$m \geq n = \mathrm{rank}A$（優決定）の場合に，もう少し複雑な前処理として，$A = QR + E$ で与えられる不完全 QR 分解の適用を検討する．ただし，$Q \in \mathbf{R}^{m\times n}$ は直交行列の近似，$R \in \mathbf{R}^{n\times n}$ は上三角行列，E は誤差行列とする．

従来はこの R を式(1.3)の前処理に用いて CG 法を適用していたが，ここでは $B = R^{-1}Q^{\mathrm{T}}$ とおいて，GMRES(k)-LS 法 1，2 を適用することを考える．

例えば GMRES(k)-LS 法 2 の場合，$R^{-1}Q^{\mathrm{T}}A\boldsymbol{x} = R^{-1}Q^{\mathrm{T}}\boldsymbol{b}$ に GMRES(k)法を適用する．ここで $R^{-1}Q^{\mathrm{T}}A$ は CGLS 法に対する式(1.3)の $\tilde{A} = R^{-\mathrm{T}}A^{\mathrm{T}}AR^{-1}$ よりも条件がよいと期待される．

3.2 IMGS(l)法

不完全 QR 分解の一つとして，修正 Gram-Schmidt 法(例えば Björck, 1996)の近似として行列 Q の現在の列ベクトルを過去の l 本の列ベクトルとのみ直交化させる，下記の IMGS(l)法を考える．

$\boldsymbol{a}_i^{(1)} = \boldsymbol{a}_i \quad (i = 1, \ldots, n)$
for $i = 1, 2, \ldots, n$
  $r_{ii} = \|\boldsymbol{a}_i^{(i)}\|_2, \quad \boldsymbol{q}_i = \dfrac{\boldsymbol{a}_i^{(i)}}{r_{ii}}$
  for $j = i+1, \ldots, \min(i+l, n)$
    $r_{ij} = \boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{a}_j^{(i)}$
    $\boldsymbol{a}_j^{(i+1)} = \boldsymbol{a}_j^{(i)} - r_{ij}\boldsymbol{q}_i$
  endfor
endfor

上記の各ステップにおいて，各 q_i は $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ の線形結合で，$\langle \boldsymbol{q}_1,\ldots,\boldsymbol{q}_n\rangle=\langle \boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\rangle$ が充たされるので，$\tilde{C}$ を正則な行列として，$Q=A\tilde{C}$ が成り立つ．従って，$Q^{\mathrm{T}}=\tilde{C}^{\mathrm{T}}A^{\mathrm{T}}$ が成り立つ．

また，$r_{ii}\neq 0\,(i=1,\ldots,n)$ より，$R=(r_{ij})$ も正則である．

従って，C を正則な行列として，$B=R^{-1}Q^{\mathrm{T}}=R^{-1}\tilde{C}^{\mathrm{T}}A^{\mathrm{T}}=CA^{\mathrm{T}}$ が成り立つ．よって，IMGS(l)法は条件 $\mathcal{R}(A)=\mathcal{R}(B^{\mathrm{T}})$ および $\mathcal{R}(A^{\mathrm{T}})=\mathcal{R}(B)$ を充たす．

ここで，下記が成り立つ．

補題 3.1.　IMGS(0)法は $B=\{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}A^{\mathrm{T}}$ とおくことに相当する．

4.　数値実験

最後にフルランクな問題に対する数値実験結果を示す．

4.1　優決定問題

まず，優決定な最小二乗問題

$$\min_{\boldsymbol{x}\in\mathbf{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2,\qquad A\in\mathbf{R}^{m\times n}\qquad(m\geq n)$$

において，係数行列 A は MATLAB 6 のコマンド sprandn で生成した．その条件数や密度(非零要素の割合)は指定し，非零要素の値は正規分布に従う乱数で発生し，非零要素のパタンも乱数を用いて定めた．具体的には，小さめの行列として，$m=1{,}000, n=320$，非零要素の密度が 4.9% の行列 RANDS2 (条件数：2×10^2)および RANDS6 (条件数：1×10^6)と，大きめの行列として，$m=10{,}000, n=1{,}000$，非零要素の密度が 1.5%で，条件数が表 1 のような行列 RANDL1–RANDL7 を用いた．

$\boldsymbol{b}$ についても同様に，各々$m=1{,}000$ および $m=10{,}000$ に対して，その各成分を正規分布に従う乱数で発生させ，$\boldsymbol{b}\notin\mathcal{R}(A)$ となるようにした．

手法としては，下記の(1)–(5)を比較した．

(1) CGLS 法：正規方程式 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ を CG 法で解く．
(2) 正規方程式に Jennings and Ajiz (1984)による IMGS 法の前処理を施してから CGLS 法を適用する．
(3) 正規方程式に提案した IMGS(l)法の前処理を施してから CGLS 法を適用する．
(4) IMGS(l)法により $B=R^{-1}Q^{\mathrm{T}}$ を求め，GMRES(k)-LS 法 1 を適用する．
(5) IMGS(l)法により $B=R^{-1}Q^{\mathrm{T}}$ を求め，GMRES(k)-LS 法 2 を適用する．

表 1.　テスト行列の条件数(優決定問題)．

Name	条件数
RANDL1	6×10^1
RANDL2	4×10^2
RANDL3	3×10^3
RANDL4	3×10^4
RANDL5	2×10^5
RANDL6	2×10^6
RANDL7	2×10^7

ただし，初期近似解は$x_0=\mathbf{0}$とし，収束の判定には，$\boldsymbol{r}=\boldsymbol{b}-A\boldsymbol{x}$を残差として，$\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2<10^{-6}$を用いた．

各手法のプログラムはMATLAB 6で書き，計算機はNECのPC（AMD Athlon XP 2000+，1.66GHz，736MB RAM）を用いた．

(2)のJenningsらのIMGS法における閾値としては$\tau=1$（正規方程式の行対角項スケーリングに相当）が一番(計算時間の上で)速かったが，下記の実験では$\tau=0.1$とおいた．Saad (1988)のIMGS法も試みたが，Jenningsらの方法よりさらに遅かった．

また，提案したIMGS(l)法においてlの値の最適値を調べるために，RANDS2，RANDS6に対して，(5)のIMGS(l)法を用いたGMRES-LS法2（リスタートなし）でlの値を変化させたところ，各々表2，表3のように$l=0$（正規方程式の行対角項スケーリングに相当）が一番速かったので下記の実験では$l=0$とおいた．これは，補題3.1にあるように，$B=CA^{\mathrm{T}}, C=\{\mathrm{diag}(A^{\mathrm{T}}A)\}^{-1}$とおくことに相当する．

さらに，表1の行列のうち，RANDL2，RANDL3に対して(IMGS(0)法を用いた)GMRES(k)-LS法2において，リスタートの周期kを変化させたところ，表4のようになった．

表4において，kはリスタート周期，iter，timeは各々収束に要する反復数と計算時間(秒)を示す．また，$*$は各問題で一番速かったものを示す．ここでは，反復数，計算時間とも，リスタートを行わない方が少なかったので，以下の実験でもリスタートなしのGMRES法を用いた．

まず，図1，2に各々上記のRANDS2，RANDS6に対する各手法の反復数vs. $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$のグラフを示す．悪条件問題RANDS6では(5)のGMRES法2は(2)の前処理CGLS法や(4)

表2. IMGS(l)法を用いたGMRES-LS法2でのlの影響(RANDS2)．収束判定：$\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2<10^{-6}$.

l	前処理時間 (秒)	反復時間 (秒)	反復数	総計算時間 (秒)
0	0.02	3.88	131	3.90
1	0.81	8.30	130	9.11
10	3.46	9.19	128	12.7
50	14.2	8.41	119	22.6
100	29.5	6.12	91	35.6
200	71.0	3.17	48	74.2
300	98.4	0.92	9	99.3
320	99.9	0.52	1	100

表3. IMGS(l)法を用いたGMRES-LS法2でのlの影響(RANDS6)．収束判定：$\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2<10^{-6}$.

l	前処理時間 (秒)	反復時間 (秒)	反復数	総計算時間 (秒)
0	0.02	19.7	320	19.7
1	0.79	30.5	320	31.3
10	3.24	32.9	320	36.1
50	13.4	32.9	320	46.3
100	27.9	33.0	320	60.9
200	68.3	15.7	192	84.0
300	97.2	1.7	24	98.9
320	97.8	0.5	1	98.3

表 4. GMRES (k)-LS **法** 2 **でのリスタート周期** k **の影響（優決定問題）．上段：リスタート周期，中段：反復数，下段: 計算時間（秒），収束判定：** $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2 < 10^{-6}$**．**

RANDL2	k	10	50	100	150	200	$\geq$ 221
	iter	402	278	238	240	223	221
	time	68.3	37.1	36.4	40.9	45.3	$*$32.5
RANDL3	k	10	50	100	200	300	$\geq$ 438
	iter	1,759	983	714	582	511	438
	time	317	174	142	129	132	$*$101

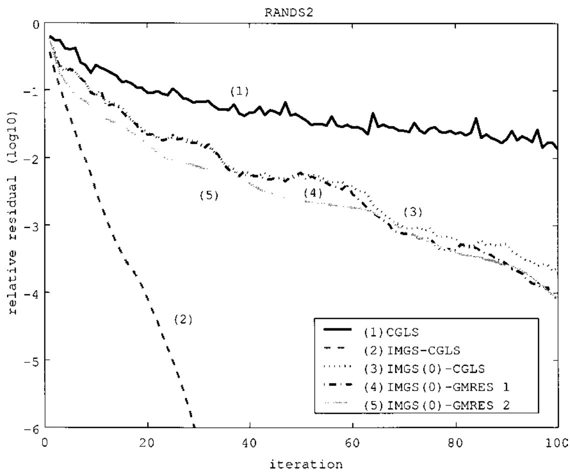

図 1. **反復数** vs. $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$ (RANDS2)**．**

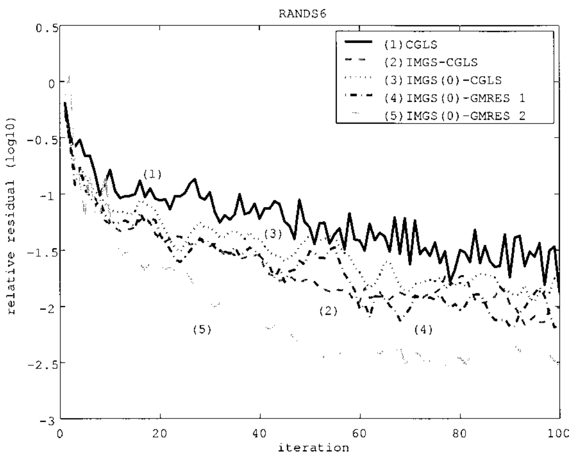

図 2. **反復数** vs. $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2$ (RANDS6)**．**

表5. 各手法の比較(優決定問題). 上段:反復数, 下段:計算時間(秒), 収束判定: $\|A^{\mathrm{T}}\boldsymbol{r}\|_2/\|A^{\mathrm{T}}\boldsymbol{b}\|_2 < 10^{-6}$.

Matrix	(1)CGLS	(2)CGLS IMGS	(3)CGLS IMGS(0)	(4)GMRES 1 IMGS(0)	(5)GMRES 2 IMGS(0)
RANDL1	131	19	58	59	57
	5.8	623	∗ 3.0	32.3	4.8
RANDL2	791	43	243	232	221
	49.0	780	∗ 15.9	308	32.5
RANDL3	5,227	109	586	458	438
	259	1,150	∗ 27.1	1,090	101
RANDL4	10,000 †	535	3,980	977	980
		2,030	∗ 207	6,630	470
RANDL5	10,000 †	884	7,901	981	971
		2,970	∗ 380	6,840	433
RANDL6	10,000 †	2,340	24,399	1,000 †	994
		5,690	1,510		∗ 462
RANDL7	10,000 †	4,998	51,560	1,000 †	983
		11,800	3,270		∗ 458

の GMRES 法1よりもよい収束性を示しているのに注意されたい.

また, 表5に表1の各問題に対する各手法の比較結果を示す. 各ますの上段は収束に要した反復数を, 下段は計算時間(秒)を示す. ただし, † はその反復数では収束しなかったことを示す.

表5から下記のことが言える.

条件数が $10^1 \sim 10^5$ の問題では(3)の IMGS(0)前処理付きの CGLS 法が一番速かった. (2)の Jennings らの IMGS 前処理付きの CGLS 法は反復数は少ないが, 計算時間がかかった. これは, 前処理や前進後退代入の部分の影響と思われる.

条件数が $10^6, 10^7$ の条件の悪い問題では, (1)の CGLS 法, (2), (3)の前処理付き CGLS 法, (4)の IMGS (0)前処理付き GMRES-LS 法1は反復数, 計算時間ともかかりすぎたのに対し, (5)の IMGS (0)前処理付きの GMRES-LS 法2は少ない反復数と計算時間で収束した. (5)が(3)の IMGS (0)前処理付きの CGLS 法より速く収束した理由としては, GMRES 法は Gram-Schmidt の直交化を陽に行うのに対し, CG 法は三項漸化式により直交化を行うため, 特に悪条件問題において GMRES 法は CG 法より丸め誤差の影響を受けにくいことが考えられる.

また, (5)の GMRES-LS 法2は(4)の GMRES-LS 法1より反復当たりの計算時間も少なくてすんでいる. これは, 方法2は $n \times n$ 行列 BA に基づくクリロフ部分空間を用いているのに対し, 方法1は $m \times m$ 行列 AB に基づくクリロフ部分空間を用いており, 今の場合 $n < m$ の優決定の最小二乗問題を扱っているためである.

4.2 劣決定問題

次に, 劣決定な最小二乗問題

$$\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2, \qquad A \in \mathbf{R}^{m\times n} \qquad (m < n)$$

に対する数値実験結果を示す. 係数行列 A は表1の各行列 RANDLn を転置したもので, 以下では RANDLnT とよぶことにする. 従って, $m = 1{,}000, n = 10{,}000$ で, 非零要素の密度は

1.5%である．また，RANDLnT の条件数は対応する元の行列 RANDLn の条件数と等しい．b は $x^* = (1, \ldots, 1)^{\mathrm{T}}$ に対して $b = Ax^*$ とおいた．

劣決定問題では $\mathrm{rank}A = m$ であれば，$Ax = b$ となる x が（無数に）存在するので収束は $r = b - Ax$ として $\|r\|_2/\|b\|_2 < 10^{-6}$ により判定した．初期近似解は $x_0 = 0$ とした．

この問題を下記の（1）–（6）の手法で解いた．

（1）CGLS 法．
（2）$C = \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ により対称に前処理した CGLS 法．（IMGS（0）法により対角項スケーリングを施した CGLS 法．）
（3）GMRES-LS 法 1 で $B = A^{\mathrm{T}}$ とおいたもの．
（4）GMRES-LS 法 1 で $B = A^{\mathrm{T}}C, C = \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ とおいたもの．
（5）GMRES-LS 法 2 で $B = A^{\mathrm{T}}$ とおいたもの．
（6）GMRES-LS 法 2 で $B = A^{\mathrm{T}}C, C = \{\mathrm{diag}(AA^{\mathrm{T}})\}^{-1}$ とおいたもの．

プログラム言語と計算機は優決定問題のものと同じである．

表 6 に各問題に対する各手法の比較結果を示す．各ますの上段は反復数を，下段は計算時間（秒）を示す．ただし，† はその反復数では収束しなかったことを示す．

表 6 より，反復数では（3）か（4）の GMRES-LS 法 1 が一番少ない反復数で収束している．計算時間では，RANDL1T から RANDL6T の問題では（1）の CGLS 法か（2）の IMGS（0）-CGLS 法が一番速く収束しているが，条件の最も悪い RANDL7T では（3）の GMRES-LS 法 1 が一番速く収束している．また，前処理をしない（1）の CGLS 法と（3）の GMRES-LS 法 1 を比べると，条件数が 10^5 以上の問題では後者の方が速く収束している．最後に，GMRES-LS 法 1 と 2 を比べると，収束までの反復数は前者の方がやや少ない上に，劣決定問題では前者の方が扱う行列が小さく，反復当たりの計算量が少ないため，速く収束している．

表 6．各手法の比較（劣決定問題）．上段：反復数，下段：計算時間（秒），収束判定：$\|r\|_2/\|b\|_2 < 10^{-6}$．

Matrix	(1) CG	(2) CG diag	(3) GMRES 1	(4) GMRES 1 diag	(5) GMRES 2	(6) GMRES2 diag
RANDL1T	127	57	113	56	115	58
	∗ 2.05	2.50	6.25	2.81	52.3	14.3
RANDL2T	773	216	428	202	439	212
	11.9	∗ 7.22	94.9	17.0	796	182
RANDL3T	4,842	538	820	429	828	453
	76.6	∗ 19.4	629	79.8	2,920	782
RANDL4T	21,818	3,855	935	974	943	981
	328	∗ 144	1,050	1,240	4,020	4,700
RANDL5T	94,806	6,683	980	976	984	982
	1,650	∗ 236	1,240	1,140	4,200	4,790
RANDL6T	138,959	21,338	943	992	950	996
	2,230	∗ 785	988	1,360	4,350	5,160
RANDL7T	179,202	34,683	872	962	882	980
	3,260	1,320	∗ 787	1,060	3,720	4,820

5. まとめ

GMRES 法を最小二乗問題に適用する方法として，もとの $m \times n$ の係数行列 A に対して，$n \times m$ の行列 B を用いて，$\min_{\boldsymbol{z} \in \mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$ に GMRES 法を適用する方法 1 と，$\min_{\boldsymbol{x} \in \mathbf{R}^n} \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$ に GMRES 法を適用する方法 2 を提案した．

また，それらの方法が任意の b に対して破綻することなく最小二乗解を与えるための行列 B に関する十分条件を導いた．さらに，B の一例として不完全 QR 分解である IMGS(l)法を提案した．

フルランクな優決定の最小二乗問題に対する数値実験によると，条件の悪い問題では IMGS(0) を用いた GMRES 法 2（正規方程式に行対角項スケーリングを施してから GMRES 法を適用することに相当する）は前処理付きの CGLS 法などの従来法よりも速く収束した．

また，フルランクな劣決定問題に対する数値実験でも，非常に条件の悪い問題では前処理なしの GMRES 法 1 が前処理付きの CGLS 法などの従来法よりも速く収束した．

謝　辞

本研究に関して有益な示唆をいただいた，張　紹良先生，杉原厚吉先生，Yimin Wei 先生，細田陽介先生，そして，多くの有益な指摘をいただいた査読者の方々に感謝いたします．また，本研究に関連する分野も含めた，幅広い分野において先駆的な研究をされてこられ，常日頃より有益な示唆をいただいている田邉國士先生に敬意と感謝を表します．本研究は文部科学省科学研究費補助金の助成を受けています．

参 考 文 献

Bai, Z.-Z., Duff, I. S. and Wathen, A. J. (2001). A class of incomplete orthogonal factorization methods. I: Methods and theories, *BIT*, **41**(1), 53–70.

Benzi, M. and Tuma, M. (2003). A robust preconditioner with low memory requirements for large sparse least squares problems, *SIAM Journal on Scientific Computing*, **25**(2), 499–512.

Björck, A. (1996). *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia.

Brown, P. N. and Walker, H. F. (1997). GMRES on (nearly) singular systems, *SIAM Journal on Matrix Analysis and Applications*, **18**(1), 37–51.

Calvetti, D., Lewis, B. and Reichel, L. (2000). GMRES-type methods for inconsistent systems, *Linear Algebra and Its Applications*, **316**, 157–169.

速水　謙(2003). GMRES and GCR(k) on singular systems, 第 32 回数値解析シンポジウム講演予稿集，2003 年 5 月，箱根，71–74.

速水　謙，伊藤徳史(2004). GMRES 法を最小二乗問題に適用するための前処理行列が充たすべき性質について，日本応用数理学会 2004 年度年会講演予稿集，212–213.

速水　謙，伊藤徳史(2005). 前処理付き GMRES 法による最小二乗問題の解法，京都大学数理解析研究所講究録 1441, 21 世紀における数値解析の新展開，2005 年 7 月，114–128.

伊藤徳史(2003). GMRES 法の線形最小二乗問題への適用，修士論文，東京大学大学院情報理工学系研究科数理情報学専攻．

伊藤徳史，速水　謙(2002). GMRES(k)法の線型最小二乗問題への適用，日本応用数理学会 2002 年度年会予稿集，http://pjsiam.jstage.jst.go.jp/ja

伊藤徳史，速水　謙(2003). 不完全 QR-GMRES(k)法による線形最小二乗問題の解法，情報処理学会第 65 回全国大会講演論文集，(1-117)–(1-118).

Ito, T. and Hayami, K. (2004). Preconditioned GMRES methods for least squares problems, NII

Tech. Report, NII-2004-006E, 1–29, National Institute of Informatics.

伊藤 徳史，速水 謙(2004)．前処理 GMRES 法による最小二乗問題の解法，2004 年度日本応用数理学会年会予稿集，210–211.

Jennings, A. and Ajiz, M. A. (1984). Incomplete methods for solving $A^{\mathrm{T}}Ax = b$, *SIAM Journal on Scientific and Statistical Computing*, **5**, 978–987.

Meijerink, J. A. and van der Vorst, H. (1977). An iterative method for linear systems of which the coefficient matrix is a symmetric M-matrix, *Mathematics of Computation*, **31**(137), 148–162.

ラオ，C. R. (1977)．『統計的推測とその応用』，東京図書，東京.

Reichel, L. and Ye, Q. (2005). Breakdown-free GMRES for singular systems, *SIAM Journal on Matrix Analysis and Applications*, **26**(4), 1001–1021.

Saad, Y. (1988). Preconditioning techniques for nonsymmetric and indefinite linear systems, *Journal of Computational and Applied Mathematics*, **24**, 89–105.

Saad, Y. and Schultz, M. H. (1986). GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing*, **7**(3), 856–869.

Tanabe, K. (1974). Characterization of linear stationary iterative processes for solving a singular system of linear equations, *Numerische Mathematik*, **22**, 349–359.

Tanabe, K. (1975). Neumann-type expansion of reflexive generalized inverses of a matrix and its hyperpower iterative method, *Linear Algebra and Its Applications*, **10**, 163–175.

張　紹良(1989)．共役残差法の一般化，博士論文，筑波大学大学院博士課程工学研究科電子情報学専攻.

Zhang, S.-L. and Oyanagi, Y. (1990). A necessary and sufficient convergence condition of orthomin(k) methods for least squares problem with weight, *Annals of the Institute of Statistical Mathematics*, **42**, 805–811.

Zhang, S.-L. and Oyanagi, Y. (1991). Orthomin(k) method for linear least squares problem, *Journal of Information Processing*, **14**, 121–125.

Solution of Least Squares Problems Using GMRES Methods

Ken Hayami[1] and Tokushi Ito[2]

[1]National Institute of Informatics
[2]Business Design Laboratory Co., Ltd.

The most commonly used iterative method for solving large sparse least squares problems $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2, A \in \mathbf{R}^{m\times n}$ is the CGLS method, which applies the (preconditioned) conjugate gradient method to the normal equation $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$.

This paper considers alternative methods by using an $n \times m$ matrix B and applying the Generalized Minimal Residual (GMRES) method, which is a robust Krylov subspace iterative method for solving systems of linear equations with nonsymmetric coefficient matrix, to $\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$ or $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$.

Next, we give a sufficeint condition concerning B for the proposed methods to give a least squares solution without breakdown for arbitrary $\boldsymbol{b}$, for over-determined, under-determined and possibly rank-deficient problems. Then, as an example for B, we propose the IMGS(l) method, which is an incomplete QR decomposition.

Finally, we show by numercial experiments on full-rank over-determined and under-determined problems that, for ill-conditioned problems, the proposed method using the IMGS(0) method, which is equivalent to diagonal scaling, gives a least squares solution faster than previous preconditioned CGLS methods.

Key words: Least squares problems, iterative method, Krylov subspace method, GMRES method, CGLS method, preconditioning.